

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP05/004653

International filing date: 16 March 2005 (16.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: JP
Number: 2004-080396
Filing date: 19 March 2004 (19.03.2004)

Date of receipt at the International Bureau: 20 May 2005 (20.05.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2004年 3月19日

出願番号
Application Number: 特願2004-080396

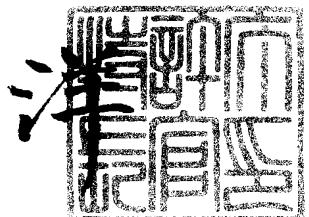
パリ条約による外国への出願に用いる優先権の主張の基礎となる出願の国コードと出願番号
The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is

出願人
Applicant(s): インターナショナル・ビジネス・マシンズ・コーポレーション

2005年 4月27日

特許庁長官
Commissioner,
Japan Patent Office

小川



【書類名】 特許願
【整理番号】 JP9040033
【提出日】 平成16年 3月19日
【あて先】 特許庁長官 殿
【国際特許分類】 G06F 13/00
G06F 9/44

【発明者】
【住所又は居所】 神奈川県大和市下鶴間 1623番地14 日本アイ・ビー・エム
株式会社 大和事業所内
【氏名】 米澤 隆

【発明者】
【住所又は居所】 神奈川県大和市下鶴間 1623番地14 日本アイ・ビー・エム
株式会社 大和事業所内
【氏名】 杉本 和敏

【発明者】
【住所又は居所】 神奈川県大和市下鶴間 1623番地14 日本アイ・ビー・エム
株式会社 大和事業所内
【氏名】 美園 和久

【特許出願人】
【識別番号】 390009531
【氏名又は名称】 インターナショナル・ビジネス・マシンズ・コーポレーション

【代理人】
【識別番号】 100086243
【弁理士】
【氏名又は名称】 坂口 博

【代理人】
【識別番号】 100091568
【弁理士】
【氏名又は名称】 市位 嘉宏

【代理人】
【識別番号】 100108501
【弁理士】
【氏名又は名称】 上野 剛史

【復代理人】
【識別番号】 100104880
【弁理士】
【氏名又は名称】 古部 次郎

【選任した復代理人】
【識別番号】 100118201
【弁理士】
【氏名又は名称】 千田 武

【選任した復代理人】
【識別番号】 100118108
【弁理士】
【氏名又は名称】 久保 洋之

【手数料の表示】
【予納台帳番号】 081504
【納付金額】 21,000円

【提出物件の目録】
【物件名】 特許請求の範囲 1

【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9706050
【包括委任状番号】 9704733
【包括委任状番号】 0207860

【書類名】特許請求の範囲

【請求項 1】

ネットワークを介して接続された複数のコンピュータによりグリッドコンピューティングを実現するコンピュータシステムにおいて、

ネットワーク上のコンピュータに対してジョブの実行要求を行うセンターサーバと、前記センターサーバからの要求に応じてジョブを実行するプロセスサーバとを備え、前記センターサーバは、

前記プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部と、

前記プロセスサーバの情報を管理すると共に、前記スケジューラ部により発行された前記リクエストを受け付けて、当該リクエストにかかるジョブを割り当てられたプロセスサーバの状況に応じて当該リクエストを当該プロセスサーバに送信するエージェント部とを備えることを特徴とするコンピュータシステム。

【請求項 2】

前記エージェント部は、複数の前記プロセスサーバに対して 1 対 1 で対応させて設けられたことを特徴とする請求項 1 に記載のコンピュータシステム。

【請求項 3】

前記エージェント部は、対応する前記プロセスサーバの能力および動作状況に関する情報を当該プロセスサーバから取得して管理し、

前記スケジューラ部は、前記エージェント部に管理されている前記情報に基づいて、前記プロセスサーバに対する前記ジョブの割り当てを行うことを特徴とする請求項 2 に記載のコンピュータシステム。

【請求項 4】

前記エージェント部は、少なくとも一部の前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信し、少なくとも他の一部の前記プロセスサーバに対して、自身が管理するタイミングで前記スケジューラ部から受け付けた前記リクエストを送信することを特徴とする請求項 1 に記載のコンピュータシステム。

【請求項 5】

少なくとも一部の前記プロセスサーバは、ファイアウォールを介して前記センターサーバに接続されており、

前記エージェント部は、前記ファイアウォールを介して接続されている前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信することを特徴とする請求項 1 に記載のコンピュータシステム。

【請求項 6】

グリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行うサーバにおいて、

システムを構成するコンピュータに対して実行すべきジョブを割り当て、当該ジョブの実行要求を行うスケジューラ部と、

前記コンピュータの情報を管理すると共に、前記スケジューラ部による前記ジョブの実行要求を当該ジョブの割り当てられた当該コンピュータに代わって受け付け、当該コンピュータの状況に応じて当該ジョブの実行要求を当該コンピュータに対して行うエージェント部とを備えることを特徴とするサーバ。

【請求項 7】

前記エージェント部は、システムを構成する前記コンピュータごとに設けられ、各々が対応するコンピュータとの間で設定された個別の通信形式で前記ジョブの実行要求を行うことを特徴とする請求項 6 に記載のサーバ。

【請求項 8】

少なくとも一部の前記エージェント部は、システムを構成する前記コンピュータからのポーリングによるアクセスに応じて当該コンピュータに対する前記ジョブの実行要求を行い、少なくとも他の一部の前記エージェント部は、自身が管理するタイミングで対応する前記コンピュータに対する前記ジョブの実行要求を行うことを特徴とする請求項7に記載のサーバ。

【請求項9】

グリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行うサーバにおいて、

システムを構成するコンピュータの能力および動作状況に関する情報を管理すると共に、当該コンピュータとの間で行われる通信を中継して、当該コンピュータの動作状況に応じて送受信を行うエージェント部と、

前記エージェント部に管理されている前記情報に基づいて前記コンピュータに実行させるジョブの割り当てを行い、ジョブを割り当てられたコンピュータに対する当該ジョブの実行要求を、当該エージェント部を介して行うスケジューラ部とを備えることを特徴とするサーバ。

【請求項10】

前記エージェント部は、システムを構成する前記コンピュータごとに設けられ、

前記スケジューラ部は、ジョブを割り当てられた前記コンピュータに対する前記エージェント部を介して当該ジョブの実行要求を行うことを特徴とする請求項9に記載のサーバ。

【請求項11】

前記スケジューラ部は、前記エージェント部に格納されている前記コンピュータの能力に関する情報に基づいて前記ジョブの割り当てを行い、かつ当該ジョブの割り当てられたコンピュータの動作状況に関わらず、当該ジョブの実行要求を行い、

前記エージェント部は、少なくとも一部の前記コンピュータに対して、当該コンピュータからのポーリングによるアクセスに応じて前記スケジューラ部によるジョブの実行要求を送信し、少なくとも他の一部の前記コンピュータに対して、自身が管理するタイミングで前記スケジューラ部によるジョブの実行要求を送信することを特徴とする請求項9に記載のサーバ。

【請求項12】

コンピュータにより、グリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行うジョブ実行制御方法であって、

前記コンピュータが、記憶装置に格納されたシステムを構成しジョブを実行するプロセスサーバの能力に基づき、当該プロセスサーバの動作状況に関わらずにジョブの割り当てを行うステップと、

前記コンピュータが、前記ジョブの割り当てられたプロセスサーバに対するジョブ実行リクエストを発行するステップと、

前記コンピュータが、発行された前記ジョブ実行リクエストを一時的に保持し、前記ジョブの割り当てられた前記プロセスサーバの動作状況に応じて、当該プロセスサーバに送信するステップと

を含むことを特徴とするジョブ実行制御方法。

【請求項13】

コンピュータに、

グリッドコンピューティングシステムを構成しジョブを実行するプロセスサーバの情報を記憶手段に格納して管理する機能と、

前記プロセスサーバの情報に基づき、当該プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行する機能と、

発行された前記リクエストを受け付けて、当該リクエストにかかるジョブを割り当てられたプロセスサーバの動作状況に応じて当該リクエストを当該プロセスサーバに送信する機能と

を実現させることを特徴とするプログラム。

【請求項 14】

前記ジョブ実行リクエストを発行する機能では、前記コンピュータに、前記プロセスサーバの動作状況に関わらず前記ジョブの割り当てを実行させることを特徴とする請求項13に記載のプログラム。

【請求項 15】

前記リクエストを前記プロセスサーバに送信する機能では、前記コンピュータに、少なくとも一部の前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記リクエストを送信させ、少なくとも他の一部の前記プロセスサーバに対して、自身が管理するタイミングで前記リクエストを送信させることを特徴とする請求項13に記載のプログラム。

【請求項 16】

前記リクエストを前記プロセスサーバに送信する機能では、前記コンピュータに、ファイアウォールを介して接続されている前記プロセスサーバに対して、当該プロセスサーバからのポーリングによるアクセスに応じて前記スケジューラ部から受け付けた前記リクエストを送信させることを特徴とする請求項13に記載のプログラム。

【書類名】明細書

【発明の名称】コンピュータシステム、これを構成するサーバ、そのジョブ実行制御方法及びプログラム

【技術分野】

【0001】

本発明は、グリッドコンピューティングに関し、特にコンピュータに対してジョブの割り当てを行うスケジューリングの方法およびこれを実現するシステム構成に関する。

【背景技術】

【0002】

近年、ネットワークで接続されたヘテロジニアス（異機種混在）な情報システムを統合して利用する、グリッドコンピューティングと呼ばれるシステムが注目されている。このシステムでは、ネットワーク上の複数のコンピュータにおけるCPUパワーやデータストレージなどのコンピュータ資源が共有され、仮想的な1つの高性能コンピュータとして利用される。複数のコンピュータに並列処理を行わせることで、1台1台の性能は低くとも高速に大量の処理を実行することが可能となる。

【0003】

グリッドコンピューティングでは、システムに与えられたジョブ（プログラムの処理単位）の実行順序を、ジョブの特性や優先順位に応じて決定するスケジューリングが非常に重要である。異機種混在環境を前提としたグリッド環境では、性能の異なる複数のコンピュータに対してジョブを割り当てる必要がある他、割り当てるジョブの大きさも様々であり、また、ローカルユーザだけでなく、遠隔地にいるユーザからもジョブを与えることから、様々なジョブの実行スケジュールを適切に設定する必要があるためである。

【0004】

グリッドコンピューティングにおけるジョブのスケジューリングの方式には、大きく分けてプッシュ（Push）タイプとプル（Pull）タイプとがある。

プッシュタイプ（例えば、非特許文献1参照）のスケジューリング方式では、スケジューラにジョブの投入があった場合に、スケジューラが、そのジョブを処理すべきコンピュータを割り当て、そのコンピュータに対してジョブの実行をリクエストする。この場合にスケジューラは、各コンピュータの利用状況をモニタリングし、ジョブの負荷情報等に応じて最適なコンピュータとジョブの割り当てを行う。

【0005】

このプッシュタイプのスケジューリング方式は、主にクラスタグリッドと呼ばれる、数百台程度のサーバ（コンピュータ）で構成され、それらを同一サイト内に配置するタイプのグリッドシステムに多く用いられる。またプッシュタイプのスケジューリング方式は、最適なコンピュータに対してSpecificにジョブを割り当てることができ、最適なスケジューリングが可能である。特に、各コンピュータの挙動にはらつきがあり、かつユーザがコンピュータを使用した場合にジョブがキャンセルされる環境では、そのような動作環境を加味した上で、最適なスケジューリングにより非常に効率の良い運用が期待できる。

【0006】

一方、プルタイプ（例えば、非特許文献2参照）のスケジューリング方式では、各コンピュータがジョブを実行可能な状態になったときにスケジューラに対してポーリング（Polling）という形でジョブのリクエストを行う。そして、このリクエストが行われた際に、実行すべきジョブがあれば、スケジューラはそのコンピュータに対してジョブを割り当てる、コンピュータが割り当てられたジョブを実行する。実行すべきジョブが無い場合は、コンピュータは、一定時間の経過後に再度ポーリングを行う。

【0007】

このプルタイプのスケジューリング方式は、非常に簡単な構成でに実装することができるため、数千台規模のグリッドシステムで多く使われる。特に、インターネット上に構築されるグリッドの場合には、ネットワークの制限のために、このプルタイプのスケジューリング方式が用いられる。またプルタイプのスケジューリング方式は、コンピュータの管

理情報が簡単で、多数のコンピュータに対して適用しやすい。そして、ポーリングによる自己最適化（余裕のあるコンピュータほど頻繁にポーリングを行う）により、ある程度の効率化は図れる。

【0008】

【非特許文献1】Chris Smith, "Open Source Metascheduling for Virtual Organizations with the Community Scheduler Framework (CSF)", Technical Whitepaper, Platform Computing Inc. 2003年8月。

【非特許文献2】Eric Korpela, Dan Werthimer, David Anderson, Jeff Cobb, Matt Lebofsky, "Massively Distributed Computing for SETI", Computing in Science & Engineering, Vol: 3, Issue: 1, Jan.-Feb. 2001, Pages: 78-83.

【発明の開示】

【発明が解決しようとする課題】

【0009】

上述したように、グリッドコンピューティングにおける従来のスケジューリング方式は、プッシュタイプの場合、最適なジョブの割り当てを行なうことができる。

しかし、最適なスケジューリングを行うために各コンピュータの使用状況を管理することが必要である。このため、スケジューラの動作が複雑になり、システム構成の変更（ジョブを実行するコンピュータの増減等）に容易に対応することができない。

また、スケジューラから各コンピュータにアクセスしてジョブの実行をリクエストする形態であるため、ファイアウォールの内側にあるコンピュータなどに対しては適用できない。

【0010】

一方、プルタイプならば、上述したように、スケジューラが必要とするコンピュータの管理情報が簡単であるため、システム構成の変更にも容易に対応でき、またコンピュータからのポーリングによってネットワーク上のアクセスが行われるため、ファイアウォールの内部にあるコンピュータに対してもジョブを割り当てることが可能である。

しかし、ポーリングによる自己最適化はあるものの、プルタイプのスケジューリング方式では、実行すべきジョブがあっても、コンピュータからのポーリングを待ってジョブの実行をリクエストするため、時間的なロスが発生してしまう。

また、ポーリングが行われた場合には、実行すべき複数のジョブのうち、そのポーリングを行ったコンピュータにとって最適なジョブを割り当てるが、そのジョブの種類によっては、より相応しい他のコンピュータがある可能性もあり、必ずしもシステム全体から見て最も効率の高い最適なスケジューリングができるわけではない。

【0011】

そこで本発明は、グリッドコンピューティングにおいて、コンピュータの種類や使用状況およびジョブの種類の両面から最適なスケジューリングを実現すると共に、システム構成の変更に対して容易に対応が可能なシステムおよびそのジョブ実行制御方法を提供することを目的とする。

また本発明は、ファイアウォール等によるアクセス制限のあるコンピュータを含むグリッドにおいても、プッシュタイプのスケジューリング方式と同等の最適なジョブの割り当てを可能とすることを他の目的とする。

また本発明は、プッシュタイプとプルタイプとを組み合わせたスケジューリング方式およびこれを用いたシステムを提供することをさらに他の目的とする。

【課題を解決するための手段】

【0012】

上記の目的を達成するため、本発明は、ネットワークを介して接続された複数のコンピュータによりグリッドコンピューティングを実現する、次のように構成されたコンピュータシステムとして実現される。すなわち、このコンピュータシステムは、ネットワーク上のコンピュータに対してジョブの実行要求を行うコンピュータであるセンターサーバと、センターサーバからの要求に応じてジョブを実行するコンピュータであるプロセスサーバ

とを備えて構成される。そして、このセンターサーバは、プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部と、プロセスサーバの情報を管理すると共に、スケジューラ部により発行されたリクエストを受け付けて、このリクエストにかかるジョブを割り当てられたプロセスサーバの状況に応じて、このリクエストをプロセスサーバに送信するエージェント部とを備える。

【0013】

ここで、より詳細には、エージェント部は、複数のプロセスサーバに対して1対1で対応させて設けられる。

また好ましくは、エージェント部は、対応するプロセスサーバの能力および動作状況に関する情報をプロセスサーバから取得して管理する。そして、スケジューラ部は、エージェント部に管理されている情報に基づいて、プロセスサーバに対するジョブの割り当てを行う。

【0014】

また、エージェント部は、プロセスサーバとのアクセスタイプに基づき、プロセスサーバからのポーリングによるアクセスに応じてスケジューラ部から受け付けたリクエストを送信したり、自身が管理するタイミングで前記スケジューラ部から受け付けたリクエストを送信したりする。

より具体的には、ファイアウォールを介してセンターサーバに接続されるプロセスサーバが存在する場合に、センターサーバからファイアウォールを越えてプロセスサーバにアクセスすることができないので、プロセスサーバからのポーリングによるアクセスを待ってリクエストを送信する。

【0015】

また、上記の目的を達成する他の本発明は、コンピュータによりグリッドコンピューティングシステムにおけるジョブのスケジューリングおよび当該ジョブの実行要求を行う、次のようなジョブ実行制御方法としても実現される。すなわち、このジョブ実行制御方法は、記憶装置に格納されたシステムを構成しジョブを実行するプロセスサーバの能力に基づき、このプロセスサーバの動作状況に関わらずにジョブの割り当てを行うステップと、ジョブの割り当てられたプロセスサーバに対するジョブ実行リクエストを発行するステップと、発行されたジョブ実行リクエストを一時的に保持し、ジョブの割り当てられたプロセスサーバの動作状況に応じて、このプロセスサーバに送信するステップとを含むことを特徴とする。

【0016】

さらに本発明は、コンピュータを制御して上述したセンターサーバの機能を実現させるプログラム、あるいはコンピュータに上記のジョブ実行制御方法の各ステップに対応する処理を実行させるプログラムとしても実現される。このプログラムは、磁気ディスクや光ディスク、半導体メモリ、その他の記録媒体に格納して配布したり、ネットワークを介して配信したりすることにより提供される。

【発明の効果】

【0017】

以上のように構成された本発明によれば、グリッドコンピューティングにおいて、スケジューラとプロセスサーバとの間の通信を中継するエージェント部をセンターサーバに設け、プロセスサーバのアクセスタイプの違いをエージェント部の制御によって吸収することにより、プロセスサーバのアクセスタイプの違いによらずに、コンピュータの種類や使用状況およびジョブの種類の両面から最適なスケジューリングを行うことができる。そしてプロセスサーバごとにエージェント部を設けることにより、プロセスサーバを増減してシステム構成を変更することができる。

【0018】

また、本発明によれば、プロセスサーバのアクセスタイプの違いをエージェント部の制御によって吸収することにより、従来はプッシュタイプのスケジューリングを行っていたジョブの実行要求を直接受け付けられるプロセスサーバと、従来はプルタイプのスケジュ

ーリングを行っていたボーリングを行った後にジョブの実行要求を受け付けるプロセスサーバとをシステム内に混在させ、かつプッシュタイプのスケジューリング方式と同等の最適なジョブの割り当てを行うことができる。

【発明を実施するための最良の形態】

【0019】

以下、添付図面を参照して、本発明を実施するための最良の形態（以下、実施形態）について詳細に説明する。

図1は、本実施形態によるグリッドコンピューティングシステムの全体構成を示す図である。

図1に示すように、本実施形態のシステムは、ジョブの割り当てを行うセンターサーバ（CS）100による割り当てにしたがって実際にジョブを実行するプロセスサーバ（PS）200とを備える。センターサーバ100とプロセスサーバ200とは、インターネットその他のコンピュータネットワークにて接続されている。このコンピュータネットワークは、通信プロトコルや、有線か、無線かといった通信形式を問わず、またファイアウォールやその他のアクセス制限を設けたものであっても良い。

【0020】

図2は、本実施形態によるグリッドコンピューティングシステムにおけるセンターサーバ100およびプロセスサーバ200を実現するのに好適なコンピュータ装置のハードウェア構成の例を模式的に示した図である。

図2に示すコンピュータ装置は、演算手段であるCPU（Central Processing Unit：中央処理装置）11と、M/B（マザーボード）チップセット12およびCPUバスを介してCPU11に接続されたメインメモリ13と、同じくM/Bチップセット12およびAGP（Accelerated Graphics Port）を介してCPU11に接続されたビデオカード14と、PCI（Peripheral Component Interconnect）バスを介してM/Bチップセット12に接続された磁気ディスク装置（HDD）15、ネットワークインターフェイス16と、さらにこのPCIバスからブリッジ回路17およびISA（Industry Standard Architecture）バスなどの低速なバスを介してM/Bチップセット12に接続されたフレキシブルディスクドライブ18およびキーボード／マウス19とを備える。

【0021】

なお、図2は本実施形態を実現するコンピュータ装置のハードウェア構成を例示するに過ぎず、本実施形態を適用可能であれば、他の種々の構成を取ることができる。例えば、ビデオカード14を設ける代わりに、ビデオメモリのみを搭載し、CPU11にてイメージデータを処理する構成としても良いし、外部記憶装置として、ATA（AT Attachment）やSCSI（Small Computer System Interface）などのインターフェイスを介してCD-R（Compact Disc Recordable）やDVD-RAM（Digital Versatile Disc Random Access Memory）のドライブを設けても良い。

【0022】

図3は、センターサーバ100の機能構成を示す図である。

センターサーバ100は、各プロセスサーバ200に対するジョブの割り当て（スケジューリング）を行うスケジューラ部110と、プロセスサーバ200の管理を行うと共にプロセスサーバ200に対するリクエストおよびレスポンスの送受信を中継するPSエージェント部120とを備える。PSエージェント部120は、グリッドコンピューティングシステムを構成するプロセスサーバ200ごとに設けられている。そして、スケジューラ部110は、このPSエージェント部120を介して各プロセスサーバ200にアクセスする。

【0023】

スケジューラ部110は、例えば図2に示したプログラム制御されたCPU11とメインメモリ13や磁気ディスク装置15等の記憶手段とで実現され、その具体的な機能として図3に示すように、PS能力検査部111と、最適PS選択部112と、ジョブ依頼部113とを備える。

P S能力検査部1 1 1は、プロセスサーバ2 0 0の能力を各プロセスサーバ2 0 0に対応するP Sエージェント部1 2 0に問合せ、情報を取得する。

最適P S選択部1 1 2は、P S能力検査部1 1 1により取得されたプロセスサーバ2 0 0の能力の情報に基づき、ジョブに応じて最適なプロセスサーバ2 0 0を選択し、割り当てを行う。このジョブの割り当てにおける最適化のロジックは任意で良い。

ジョブ依頼部1 1 3は、最適P S選択部1 1 2において選択されたプロセスサーバ2 0 0のP Sエージェント部1 2 0に対してジョブ実行を要求するリクエストを発行する。

【0 0 2 4】

本実施形態では、P Sエージェント部1 2 0がスケジューラ部1 1 0とプロセスサーバ2 0 0との間の通信を中継し、スケジューラ部1 1 0によるジョブの実行要求をプロセスサーバ2 0 0に代わって受け付ける。そのため、P S能力検査部1 1 1の問い合わせ先やジョブ依頼部1 1 3のリクエストの発行先がP Sエージェント部1 2 0となっているが、上述したように、スケジューラ部1 1 0の機能自体は、既存のプッシュタイプのスケジューラと変わらない。したがって、スケジューラ部1 1 0には、既存のグリッドコンピューティングシステムで利用されているスケジューラを用いることができる。

【0 0 2 5】

P Sエージェント部1 2 0は、例えば図2に示したプログラム制御されたC P U1 1とメインメモリ1 3や磁気ディスク装置1 5等の記憶手段とで実現され、その具体的な機能として図3に示すように、P S状況管理部1 2 1と、P S能力管理部1 2 2と、ジョブ受け付け部1 2 3と、ジョブ依頼部1 2 4と、ポーリング待ち受け部1 2 5とを備える。

P S状況管理部1 2 1は、対応するプロセスサーバ2 0 0とアクセスして、プロセスサーバ2 0 0の現在の動作状況を把握する。

P S能力管理部1 2 2は、プロセスサーバ2 0 0のジョブ実行能力に関する統計情報等を管理し、スケジューラ部1 1 0のP S能力検査部1 1 1からの問い合わせに応じて、管理している情報を返す。ここで、ジョブ実行能力に関する統計情報等とは、単にC P U自身の処理能力や記憶装置の記憶容量の静的な情報だけでなく、C P Uに対する負荷の時間変動や動作傾向といった動的な内容を統計的に処理して得られた情報を含む。

P S状況管理部1 2 1およびP S能力管理部1 2 2に管理される情報は、P Sエージェント部1 2 0が対応しているプロセスサーバ2 0 0から取得され、図2のメインメモリ1 3や磁気ディスク装置1 5等の記憶手段に格納される。

ジョブ受け付け部1 2 3は、スケジューラ部1 1 0のジョブ依頼部1 1 3から発行されたジョブ実行リクエストを受け付ける。

ジョブ依頼部1 2 4は、ジョブ受け付け部1 2 3により受け付けられたジョブ実行リクエストを対応するプロセスサーバ2 0 0に送信する。

ポーリング待ち受け部1 2 5は、プロセスサーバ2 0 0からのポーリングによる、ジョブの実行が可能であるという通知を受付ける。

【0 0 2 6】

詳しくは後述するが、本実施形態では、プッシュタイプのスケジューラを用いながら、プロセスサーバ2 0 0からのポーリングによるアクセスを受け付けてジョブ実行リクエストを送信することができる。P Sエージェント部1 2 0のポーリング待ち受け部1 2 5は、このポーリングの受け付けのために用いられるものであり、ポーリングを行わずにセンターサーバ1 0 0における所望のタイミングで送信されたジョブ実行リクエストを受け付けられるプロセスサーバ2 0 0のP Sエージェント部1 2 0に対しては、必要な構成要素ではない。

【0 0 2 7】

次に、プロセスサーバ2 0 0の機能構成と、対応するP Sエージェント部1 2 0との関係について説明する。

上述したように、本実施形態において想定されるプロセスサーバ2 0 0は、ポーリングを行わないものと行うものとがある。

図4は、ポーリングを行なうプロセスサーバ2 0 0の機能構成とP Sエージェント部1 2

0との関係を示す図である。

プロセスサーバ200は、図2に示したようなコンピュータ装置をグリッドコンピューティングシステムにおけるプロセスサーバ200として機能させるためのPS組み込み部210を備えている。

【0028】

PS組み込み部210は、例えば図2に示したプログラム制御されたCPU11とメインメモリ13や磁気ディスク装置15等の記憶手段とで実現され、その具体的な機能として図4に示すように、PS状況監視部211と、PS状況通知部212と、ジョブ受け付け部213と、ジョブ実行部214と、ポーリング実行部215とを備える。

PS状況監視部211は、自装置（プロセスサーバ200）の現在の使用状況や資源の状況を監視し、情報を収集する。

PS状況通知部212は、PS状況監視部211において収集されたPSの使用状況や資源の状況に関する情報を、センターサーバ100のPSエージェント部120に通知する。PSエージェント部120では、PS状況管理部121およびPS能力管理部122がこの通知を受け付け、メインメモリ13や磁気ディスク装置15等の記憶装置に格納して管理する。PS状況通知部212からセンターサーバ100への通知は、定期的に行っても良いし、プロセスサーバ200の動作状況が変更された際に行うようにしても良い。また、センターサーバ100の対応するPSエージェント部120から任意のタイミングでプロセスサーバ200へ問い合わせても良い。

ジョブ受け付け部213は、センターサーバ100のPSエージェント部120のジョブ依頼部124から送信されたジョブ実行リクエストを受け付ける。

ジョブ実行部214は、ジョブ受け付け部213により受け付けられたジョブを、プロセスサーバ200の資源を用いて実行する。

ポーリング実行部215は、PS状況監視部211の監視等に基づき、プロセスサーバ200がジョブを実行可能な状況である場合に、その旨を示す通知をセンターサーバ100のPSエージェント部120に対して行う。PSエージェント部120では、ポーリング待ち受け部125がポーリング実行部215からの通知を受け付けて、ジョブ依頼部124にジョブ実行リクエストの送信を行わせる。

【0029】

なお、ポーリング実行部215は、プロセスサーバ200がポーリングを行わずにセンターサーバ100から送信されたジョブ実行リクエストを受け付けられる場合、すなわちファイアウォールの内側にある等のアクセス制限がない場合は、設ける必要はない。

図5は、ポーリングを行わないプロセスサーバ200の機能構成とPSエージェント部120との関係を示す図である。

【0030】

次に、上記のように構成された本実施形態のグリッドコンピューティングシステムの動作について説明する。

上述したように、センターサーバ100のスケジューラ部110は、既存のプッシュタイプのスケジューラと同様であるので、実行すべきジョブが発生すると、センターサーバ100の管理下にある全てのプロセスサーバ200を対象として、当該ジョブの割り当てを行う。このとき、PSエージェント部120からプロセスサーバ200の能力や動作傾向等の統計情報等を取得し、これらの情報およびジョブの種類や特性に基づいて最適なスケジューリングを行う。そして、ジョブが割り当てられたプロセスサーバ200の動作状況に関わらずにジョブ実行リクエストを発行し、そのプロセスサーバ200に対応するPSエージェント部120に送信してジョブの実行要求を行う。

【0031】

PSエージェント部120の動作は、プロセスサーバ200のアクセスタイプが、センターサーバ100からのジョブの実行要求を直接受け付けられるタイプか、ファイアウォールの内側等にありポーリングを行った後にジョブの実行要求を受け付けるタイプかによって、動作が異なる。

図6は、各PSエージェント部120によるジョブの実行要求時の動作を説明するフローチャートである。

図6に示すように、PSエージェント部120は、ジョブ受け付け部123によりスケジューラ部110からジョブ実行リクエストを受け付け（ステップ601）、当該PSエージェント部120に対応するプロセスサーバ200がジョブの実行要求を直接受け付けられるタイプならば、そのジョブ実行リクエストを直ちにプロセスサーバ200へ送信する（ステップ602、604）。

【0032】

一方、当該PSエージェント部120に対応するプロセスサーバ200がポーリングを行った後にジョブの実行要求を受け付けるタイプならば、PSエージェント部120は、プロセスサーバ200によるポーリングが行われるまで待機し（ステップ602、603）、ポーリングが行われた後にスケジューラ部110から受け付けたジョブ実行リクエストをプロセスサーバ200へ送信する（ステップ604）。なお、PSエージェント部120は、プロセスサーバ200によるポーリングが行われた際に送信すべきジョブ実行リクエストを持っていない場合は、何ら処理を行わず、ジョブの受け付けおよび次のポーリングのタイミングを待ち受ける。

【0033】

プロセスサーバ200は、センターサーバ100からのジョブの実行要求を直接受け付けられるタイプである場合、センターサーバ100のPSエージェント部120からジョブ実行リクエストを受け付けると、そのリクエストに応じてジョブを実行し、実行結果をセンターサーバ100のPSエージェント部120に返す。

一方、プロセスサーバ200がポーリングを行った後にジョブの実行要求を受け付けるタイプである場合、そのプロセスサーバ200は、ジョブの実行が可能な動作状況となつた時点でポーリングを行い、ジョブ実行リクエストが送られるのを待つ。そして、センターサーバ100における対応するPSエージェント部120から送信されたジョブ実行リクエストを受け付けたならば、そのリクエストに応じてジョブを実行し、実行結果をセンターサーバ100のPSエージェント部120に返す。ジョブ実行リクエストを受信しなかつた場合は、一定時間経過後に再度ポーリングを繰り返す。

【0034】

以上のプロセスサーバ200の動作は、従来のグリッドコンピューティングシステムにおけるプロセスサーバの動作と同様である。ただし、上述したように本実施形態では、センターサーバ100において各プロセスサーバ200に対応するPSエージェント部120が、プロセスサーバ200のアクセスタイルに応じて、自身が管理するタイミングでジョブの実行要求を行うか、またはプロセスサーバ200からのポーリングによるアクセスを待つでジョブの実行要求を行うかを制御できる。すなわち、プロセスサーバ200のアクセスタイルの違いをPSエージェント部120の制御によって吸収できるため、本実施形態によるシステムは、センターサーバ100からのジョブの実行要求を直接受け付けられるプロセスサーバ200と、ファイアウォールの内側等にありポーリングを行った後にジョブの実行要求を受け付けるプロセスサーバ200とを混在させることができる。

【0035】

また、本実施形態は、上記のようにポーリングを行った後にジョブの実行要求を受け付けるプロセスサーバ200を含む構成とすることができます。そして、この場合でも、スケジューラ部110は、PSエージェント部120を介してプロセスサーバ200に対するジョブの実行要求を行うので、プロセスサーバ200の動作状況の影響を受けずに、プロセスサーバ200の能力とジョブの種類や特性とに基づく最適なスケジューリングを行うことができる。

【0036】

さらに本実施形態は、上述したように、プロセスサーバ200ごとに1対1で対応するPSエージェント部120をセンターサーバ100に設け、このPSエージェント部120に、対応するプロセスサーバ200に関する情報の管理およびリクエストやレスポンス

の送受信の制御を行わせる。また、スケジューラ部110は、PSエージェント部120が管理するプロセスサーバ200の情報に基づいて、プロセスサーバ200に対するジョブの割り当てを行う。このため、プロセスサーバ200を増減してシステム構成を変更する場合にも、対応するPSエージェント部120を追加または削除することによって、容易に対応することができる。

【図面の簡単な説明】

【0037】

【図1】本実施形態によるグリッドコンピューティングシステムの全体構成を示す図である。

【図2】本実施形態によるグリッドコンピューティングシステムにおけるセンターサーバおよびプロセスサーバを実現するのに好適なコンピュータ装置のハードウェア構成の模式的に示した図である。

【図3】本実施形態におけるセンターサーバの機能構成を示す図である。

【図4】本実施形態におけるポーリングを行うプロセスサーバの機能構成とPSエージェント部との関係を示す図である。

【図5】本実施形態におけるポーリングを行わないプロセスサーバの機能構成とPSエージェント部との関係を示す図である。

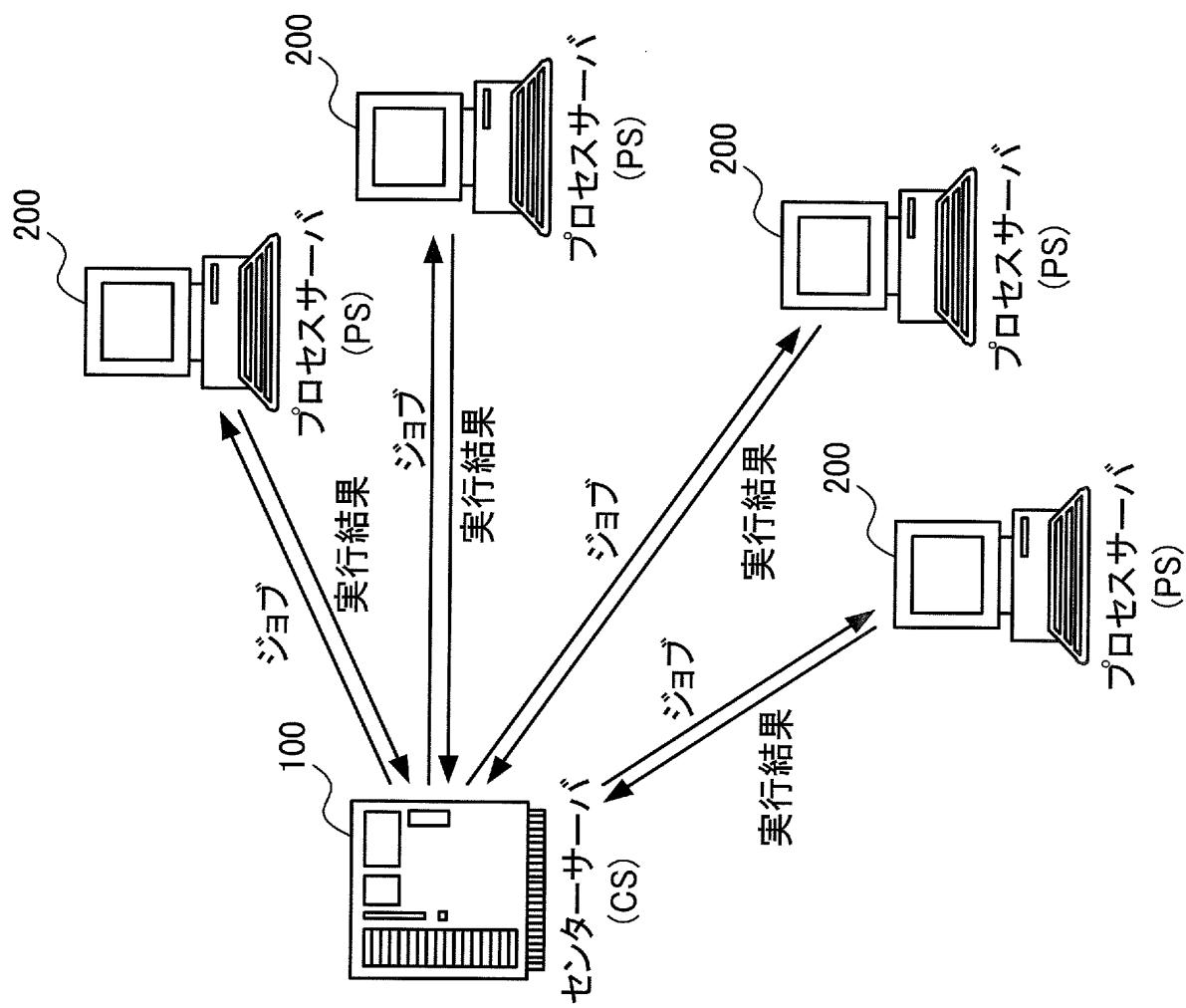
【図6】本実施形態における各PSエージェント部によるジョブの実行要求時の動作を説明するフローチャートである。

【符号の説明】

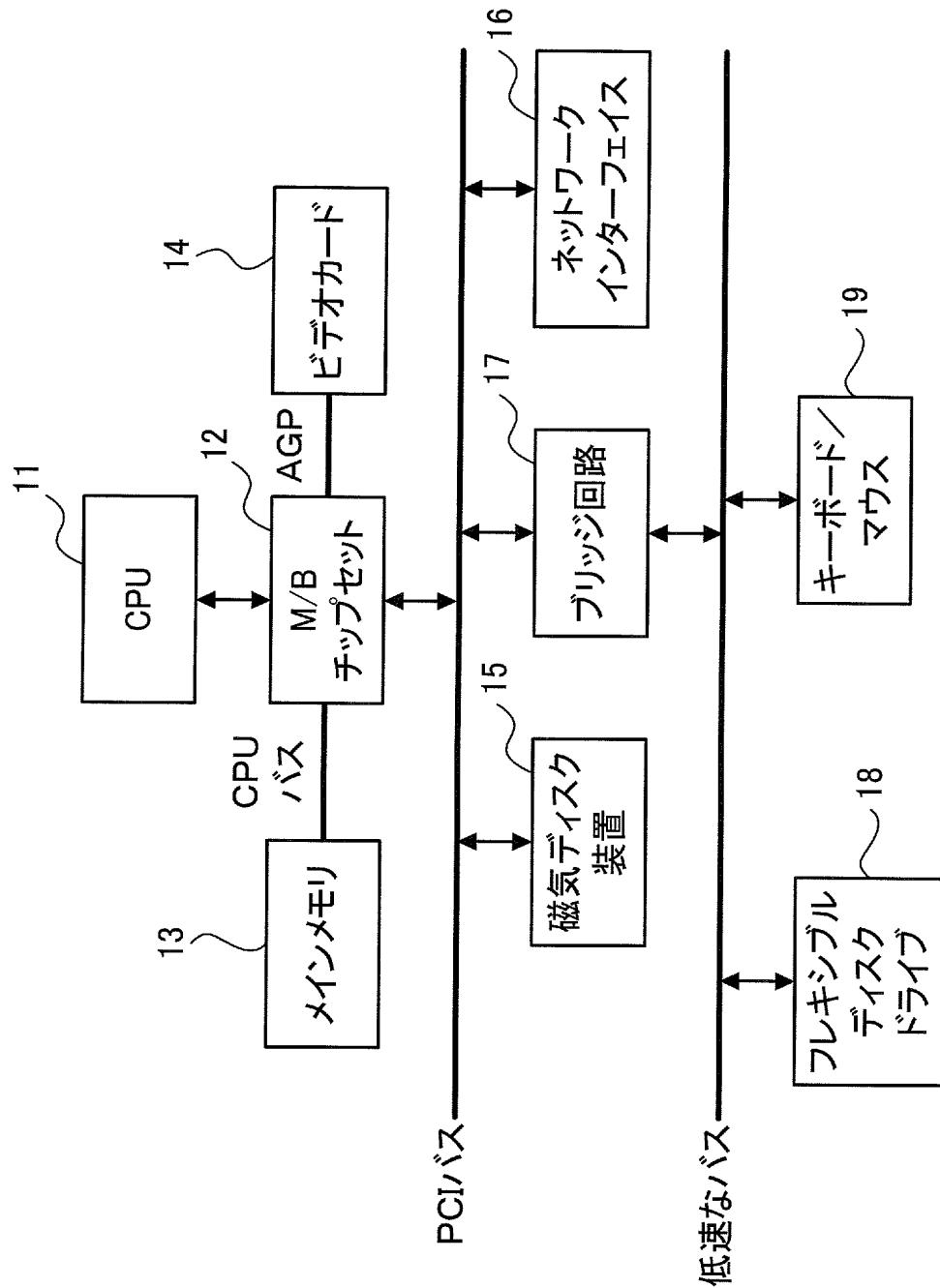
【0038】

11…CPU (Central Processing Unit: 中央処理装置)、13…メインメモリ、15…磁気ディスク装置 (HDD)、100…センターサーバ、110…スケジューラ部、111…PS能力検査部、112…最適PS選択部、113…ジョブ依頼部、120…PSエージェント部、121…PS状況管理部、122…PS能力管理部、123…ジョブ受け付け部、124…ジョブ依頼部、125…ポーリング待ち受け部、200…プロセスサーバ、210…PS組み込み部、211…PS状況監視部、212…PS状況通知部、213…ジョブ受け付け部、214…ジョブ実行部、215…ポーリング実行部

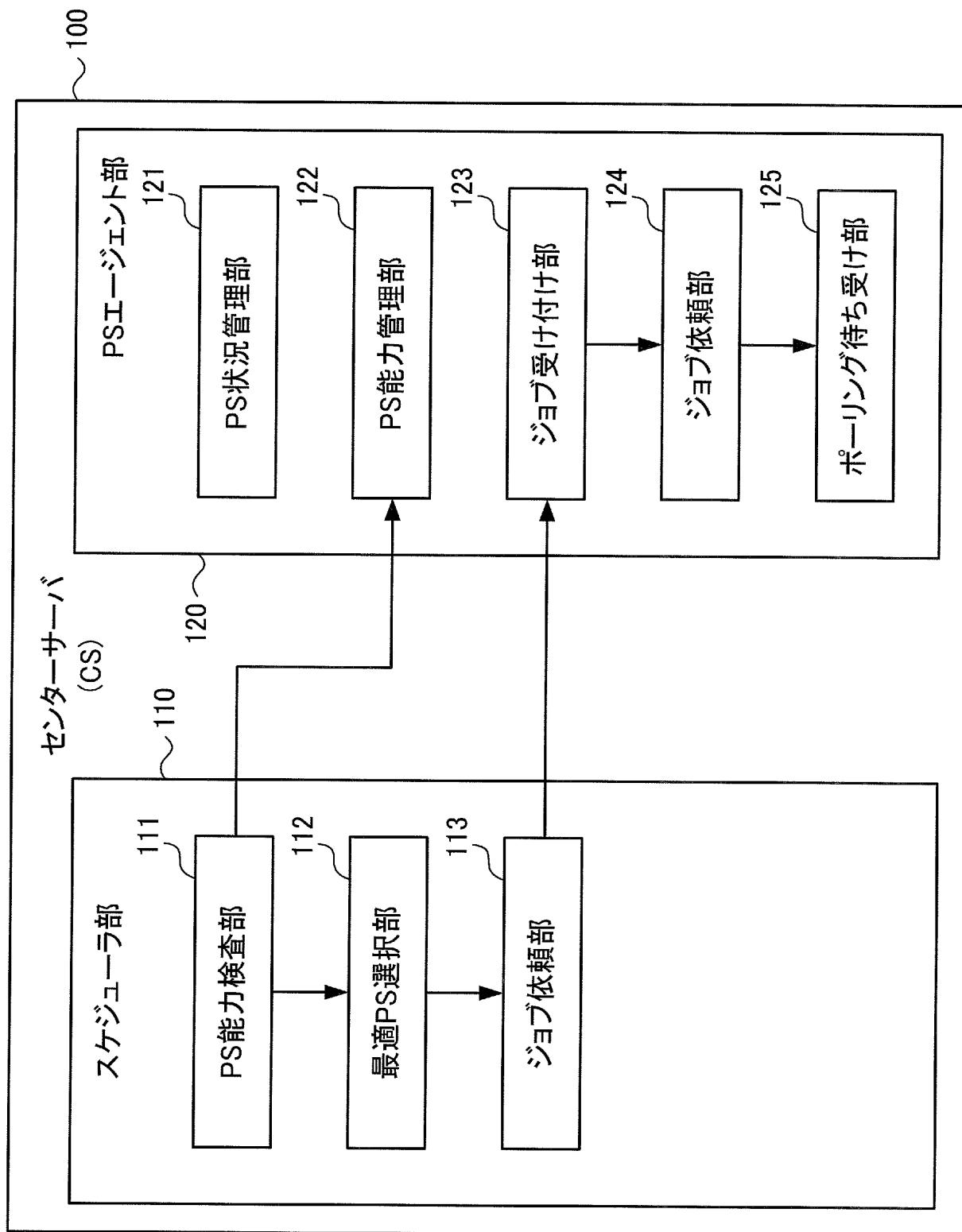
【書類名】図面
【図 1】



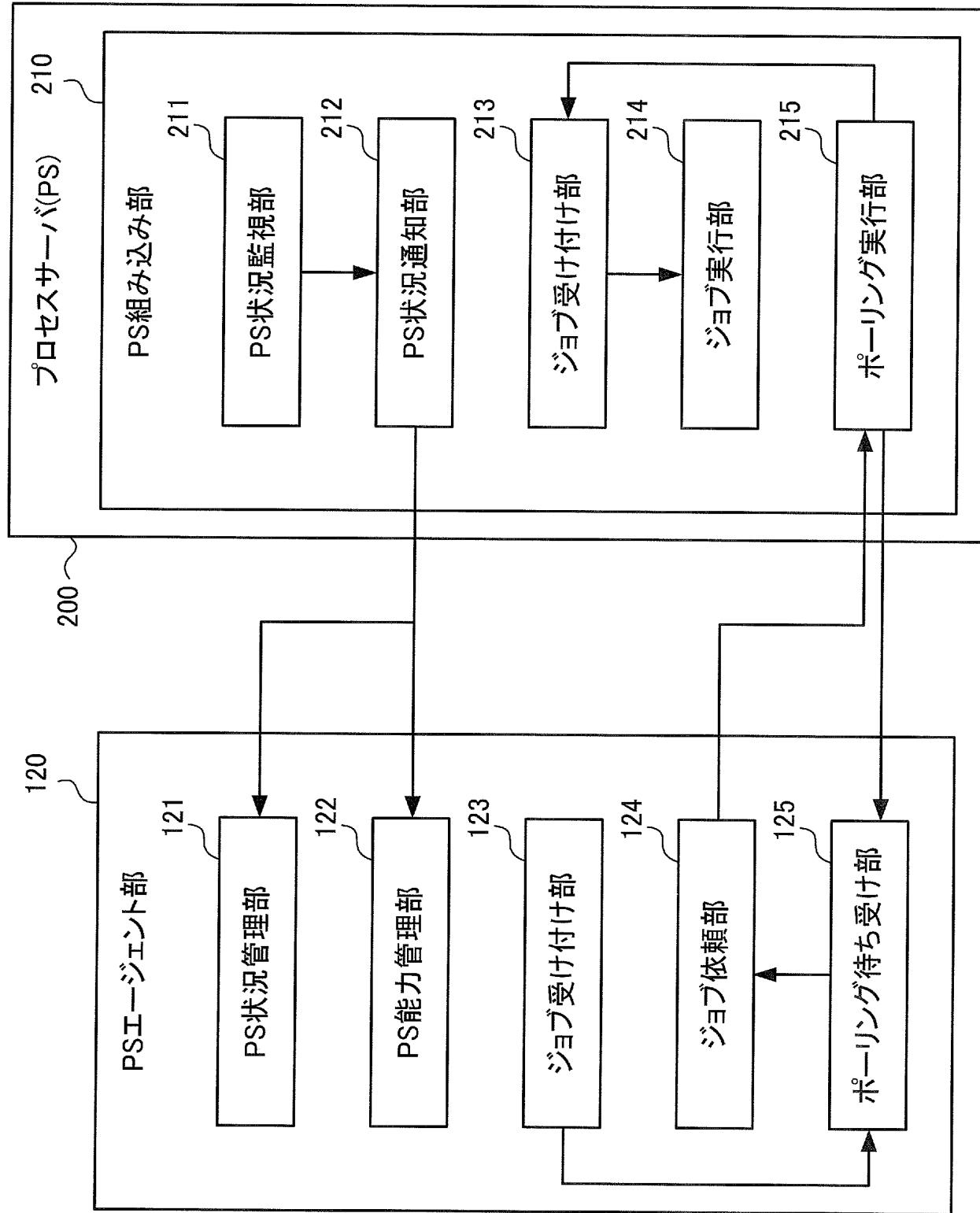
【図 2】



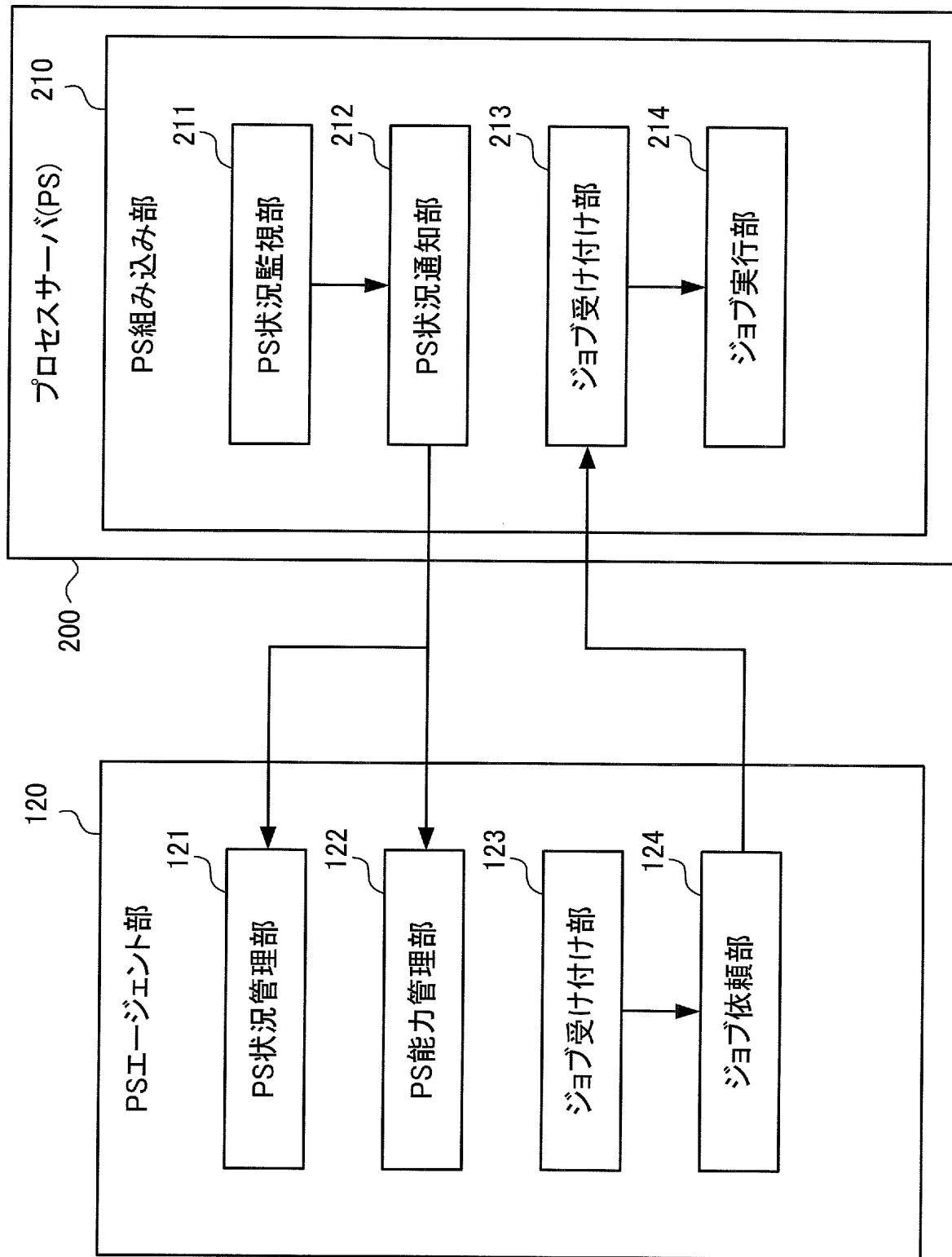
【図3】



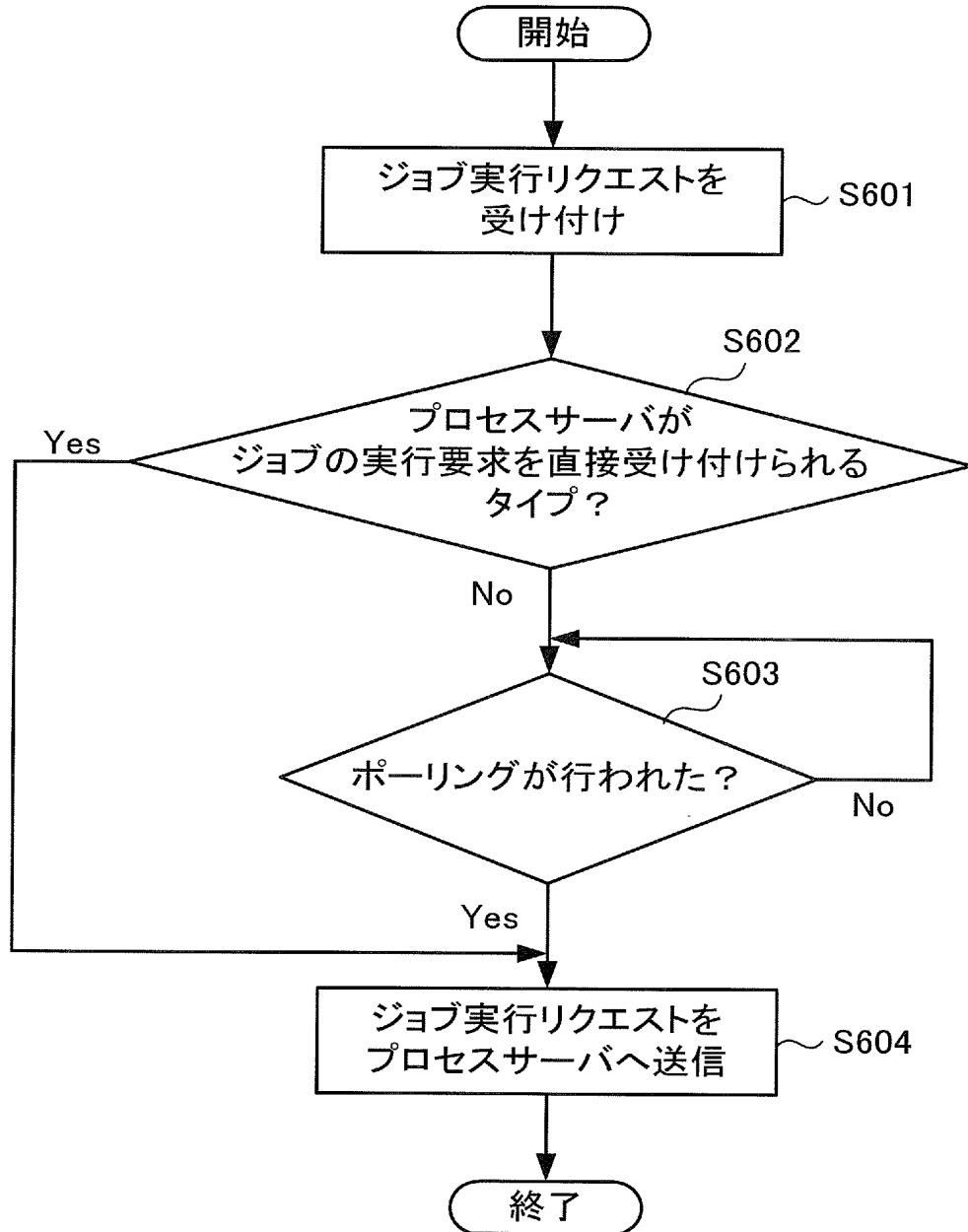
【図 4】



【図 5】



【図 6】



【書類名】要約書

【要約】

【課題】 グリッドコンピューティングにおいて、コンピュータの種類や使用状況およびジョブの種類の両面から最適なスケジューリングが可能であり、システム構成の変更に対して容易に対応が可能なシステムを実現する。

【解決手段】 グリッドコンピューティングにおいてネットワーク上のプロセスサーバに対してジョブの実行要求を行うセンターサーバ100が、プロセスサーバに対して実行すべきジョブを割り当て、ジョブ実行リクエストを発行するスケジューラ部110と、プロセスサーバ200の情報を管理すると共に、スケジューラ部110により発行されたリクエストを受け付けて、このリクエストにかかるジョブを割り当てられたプロセスサーバ200の動作状況やアクセスタイプに応じて、このリクエストをプロセスサーバ200に送信するPSEージェント部120とを備える。

【選択図】 図3

出願人履歴

3 9 0 0 0 9 5 3 1

20031205

住所変更

5 9 1 0 8 8 7 9 9

アメリカ合衆国 1 0 5 0 4 ニューヨーク州 アーモンク ニュ
ー オーチャード ロード
インターナショナル・ビジネス・マシンズ・コーポレーション